DOCUMENT RESUME

ED 049 301 TM 000 490

AUTHOR Owens, William A.
TITLE A Quasi Actuarial Prospect for Individual Assessment.
PUB DATE [70]
NOTE 16p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Behavioral Science Research, Behavior Patterns, *Biographical Inventories, Cluster Grouping, Employee Attitudes, Employment Potential, Experience, Group Behavior, *Individual Characteristics, *Industrial Personnel, *Models, Personality Assessment, Personnel Data, *Predictive Measurement, Predictor Variables, Psychology, Recruitment, Research Methodology

ABSTRACT

A conceptual model of individual assessment through the use of biodata responses with minimal input information is outlined. The process is considered especially applicable to industrial psychology. A scored autobiographical data form, which measures the individual's past behavior and experiences, provides for assignment to a specific subgroup characterized by a cluster of distinctive behavioral actions as determined by marker variables. It is expected that the use of such subgroups for additional experiments and field studies in psychological and behavioral science research will in turn provide more feedback definition of the subgroups. The factoral dimensions of the subgroup profiles include: academic achievement, tough-mindedness, goals, authoritarianism, and maladjustment. Assumptions emerging from the study are that past behavior is a good predictor of concurrent and future behavior, and confidence in post hoc validity of biodata-based subgrouping appears justified. The model's thesis is that distinctive, differental behavior can be associated with subgroup membership and by matching an individual's biodata profile with the subgroup profile modal behavior for the individual can be predicted. Present performance, promotability, sociability, motivation, managerial style, individual responsibility, innovativeness are seen as predictor variables in employee behavior through implementation of the model. Advantages and disadvantages of the model are also included. Statistical data on subgroup validity is presented. (AE)

ED049301

A QUASI ACTUARIAL PROSPECT FOR INDIVIDUAL ASSESSMENT

William A. Owens

University of Georgia

The term assessment has been used in contexts so varied that Fiske and Pearson (1970) in their recent chapter for the Annual Review say they are dropping it because of its imprecision. It has, in fact, been used on the one hand to refer to the combined effect of several cognitive predictors of overall performance expressed via multiple R; and on the other, it has also been applied to the inspection and judgmental summarization of projective test protocols with the ultimate goal of achieving a "good description" of the S and with only incidental criterion implications, if any. Thus, so-called assessment has run the full gamut from "actuarial" to "clinical." Whereas, most of us in Industrial Psychology probably feel more comfortable with the actuarial variety, we probably also would admit a gnawing feeling that our characterizations of a S ought to be fuller, more complete and more meaningful if only such embellishments could be achieved without an accompanying loss in validity. In short, we may be haunted by a sophomoric feeling that a psychologist should be a man capable of responding convincingly to the question, "what kind of person is John Jones?"

An overlapping implication in the use of the term assessment is that the resulting appraisal is exhaustive. Undiluted prediction of a present criterion does not require this, but description, comprehension, and the potential prediction of future criteria, as yet unknown, all recommend it. The classic Michigan study of Kelley and Fiske (1950) required that each

TM 000 490

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

S invest a week. This may well be one limiting case; but, if it is, the other end of the scale is anchored by that of the job applicant who responds to a 15' mental ability test and winds up with a complete characterization consisting of one score.

The point of these comments is that during the next few minutes I hope to suggest the broad outlines of an assessment scheme capable of providing both actuarial and descriptive data; one capable of providing great amounts of information about the individual without requiring great amounts of time from him, and one which I believe offers much for the future of industrial psychology.

At the University of Georgia my colleagues, Lyle Schoenfeldt, *Patrick Pinto, and I have been involved with the evaluation of a conceptual model for biodata research (Owens, 1968). As a spin-off, this investigation seems to us to offer a new look at the problem of individual assessment. A hasty review of the basic model will provide insight into its application.

In substance, it is an integrative model designed to draw measurement and experimentation, and indeed all the behavioral sciences, somewhat closer together. For the student of measurement and its applications it provides the antecedent information required for "causal-type inferences" and enhanced understanding. For the experimentalist it offers the possibility of employing relatively more homogeneous subsets of Ss, and thus of reducing error; it also affords knowledge of the Ss pre-experimental experience and behavior. For the behavioral scientist, generally, it provides a conceptual framework to which anyone who systematically observes human behavior may attach his findings, with the prospect that they will

---

* Now at I. R. C., The University of Minnesota.

both gain a new perspective and enhance the meaningfulness of the system.

A sort of flow chart representation of the basic model appears in figure 1. Here the large circle on the left represents a pool of Ss, let us say 850 male college freshmen, all of whom complete a scored autobiographical data form designed to tap the significant dimensions of their prior experiences and behaviors. Biodata, known to be a fine empirical predictor, directly implements the basic measurement axiom that the best such predictor of man's future behavior is his past behavior. Our concern, however, is with a corollary; namely that, if the axiom holds, it should then follow that Ss who have behaved similarly in the past will continue to behave similarly in the future.

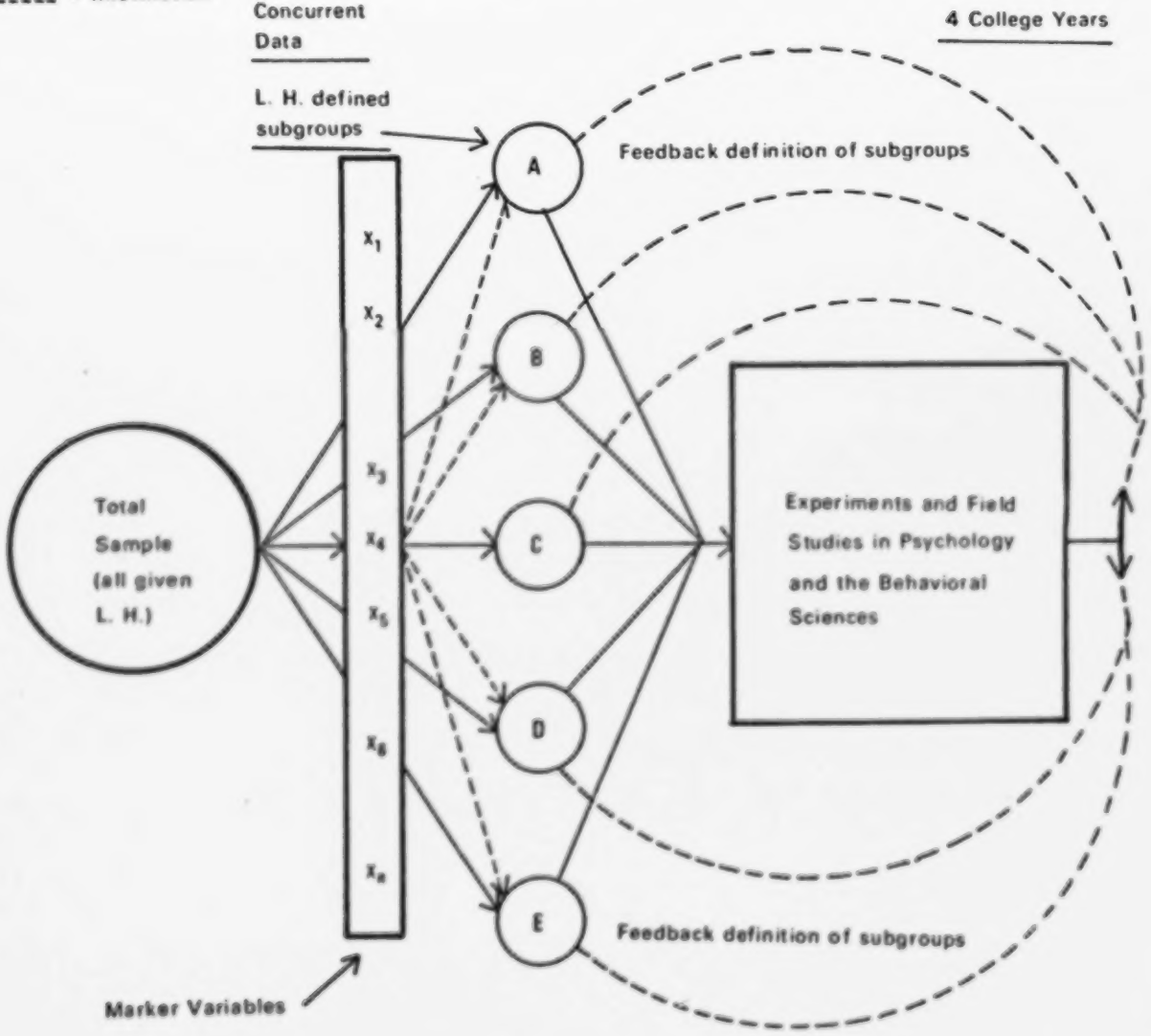
Thus, what we wish to do is to sort our Ss into subsets (here represented as A through E) such that these subsets will display internally similar and externally differing patterns of prior experience. To achieve greater brevity, let us deal with the operations involved at a conceptual level only. An excellent and detailed discussion on methodology is available elsewhere (See Schoenfeldt, 1970). What we have done to identify Ss with comparable patterns of prior experience is to factor their biodata responses; to profile each S on the resulting dimensions; to obtain a matrix of the distances between each profile and each other; to hierarchically group the profiles into "families" according to the method described by Ward and Hook (1963); and finally, to "clean up" the assignments of Ss to these subgroups in several ways recommended by the approximate nature of the methods employed to identify the groups. Ultimately, then, persons within a subgroup have similar patterns of prior experience, as expressed in similar biodata profiles, and should thus tend, by hypothesis, to exhibit similar patterns of concurrent and future behavior.

Historic
Data

____ = Ss

\- - - - - = Information

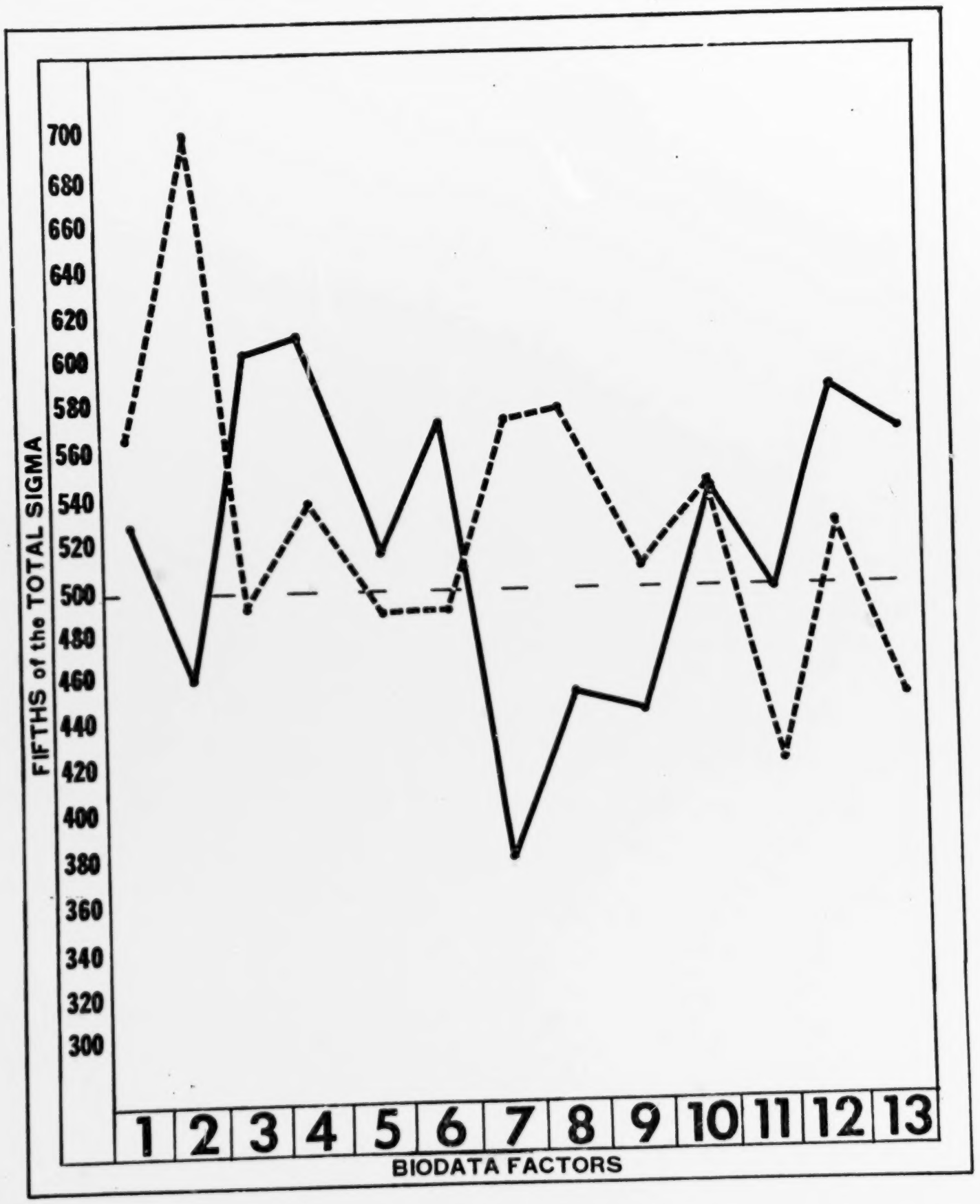
DEVELOPMENTAL - INTEGRATIVE MODEL



What follows is broadly concerned with the significance and meaningfulness of this subgrouping; i.e., what are the behavioral correlates of subgroup membership, and does such membership <u>really</u> argue for distinctive behavior. Thus, X_1 X_2--- X_n are a hopefully representative spectrum of "reference measures" or "marker variables" to be employed in characterizing the subsets, but also useful in attesting to whether or not they really <u>differ</u> in terms of concurrent performance. Even more crucially, the box on the right, were it "opened," could be seen to contain a series of experiments and field studies designed to establish whether or not subgroups selected with an expectation of differential performance actually <u>do</u> behave differently. Finally, the dotted lines returning to the subgroups A through E indicate continuous feedback regarding differential subgroup behaviors such that these may be collated and cataloged to provide definitions of ever increasing meaning and precision.

Optimistically, all of this may sound both entertaining and plausible, yet conceptual models are easy to draw and hard to validate. The crucial question remains <u>what does it mean to belong to a given subgroup</u>? The query can probably best be answered by comparing the test behaviors, field study records and experimental results characteristic of several specific subgroups. Accordingly, let us turn next to figure 2.

Appearing here are the biodata profiles of our 23 male subgroups. The upper one, #9, contains 29 Ss; and the lower one, #21, contains 22 Ss. These groups were chosen because they differ, although they are <u>not</u> the most different. The biodata dimensions of their definitions carry consensus labels as follows: parental warmth, intellectualism, academic achievement, social introversion, scientific interest, socioeconomic status, independence-dominance, parental control, academic attitude, sibling friction, religious


FIFTHS of the TOTAL SIGMA
700
680
660
640
620
600
580
560
540
520
500
480
460
440
420
400
380
360
340
320
300
1 2 3 4 5 6 7 8 9 10 11 12 13
BIODATA FACTORS

activity, athletic interest and social desirability. Each of the 23 subgroups differs significantly from a number of the others on each of these 13 dimensions--a less than remarkable fact since these were the vehicle for the subgrouping. Specifically, subgroups 9 and 21 differ significantly on 7 factors, and their means sometimes differ by as much as two sigma of the total sample (units at left = 1/5 σ). This slide was included for two reasons. First, it suggests something of the basis for and effectiveness of the subgrouping; and second, it shows what a subgroup really <u>is</u>; namely, a distinctive profile--or band of values--across 13 biodata defined dimensions of prior experience. If we accept the fact of readily apparent subset differences at this stage, the crucial question remains, "are there significant and differential behavioral correlates of belonging to a subgroup?"

One sort of answer may be obtained by examining the distinctiveness of subgroup performance on the marker or reference variables referred to in slide #1. Accordingly, in figure 3 you may observe the marker variable profiles of the biodata defined subgroups 9 and 21.

The factored dimensions of these profiles and the scales loading on them are:

<u>Academic Achievement</u>

H.S./G.P.A.

SAT-V & Q

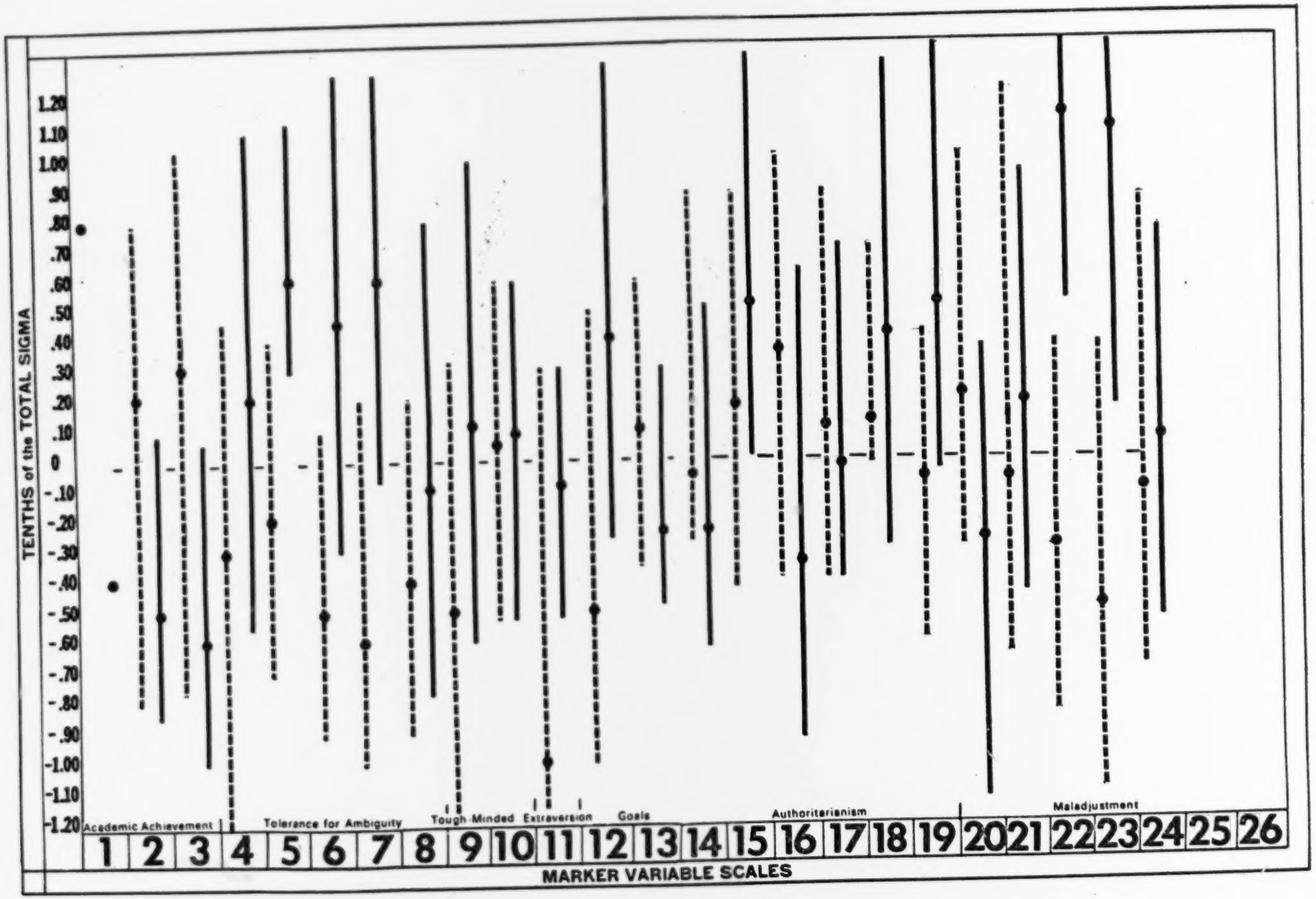
<u>Tolerance for</u> Ambiguity

cognitive values

positive emotionality

emotional exposure

integrative complexity and hierarchical complexity


TENTHS of the TOTAL SIGMA
1.20
1.10
1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
0
-.10
-.20
-.30
-.40
-.50
-.60
-.70
-.80
-.90
-1.00
-1.10
-1.20
Academic Achievement
Tolerance for Ambiguity
Tough-Minded
Extraversion
Goals
Authoritarianism
Maladjustment
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
MARKER VARIABLE SCALES

Tough-Mindedness

tough vs. tender minded

emotional inhibition

Goals

short term goals

long term goals

Authoritarianism

radicalism vs. conservatism

direct and reverse F

social-religious conformity

physical goals and

economic goals

Maladjustment

social desirability

externalization vs. internalization

negative emotionality

conceptual simplicity and

neuroticism

The circle or dot indicates the subset mean, the length of the bar represents the interquartile range, and the units are tenths of a sigma for the total sample. Overall the entire 23 subgroups differ significantly on all but 2 of these 24 variables and the two subgroups shown differ significantly ($p=<.01$) on 5 of them. Differences are also both large and significant across the occupational scales of the Strong Vocational Interest Blank not shown. Indeed, score on most V.I.B. scales can be very well predicted from biodata antecedents. The point of the moment, however, is only that biodata defined subsets of persons do differ substantially

in their performances across a broad spectrum of concurrent measures. Hopefully, this evidence is convincing, but the skeptic might still say, "Of course one paper and pencil performance will predict another"; and the outright cynic would suspect that the same items, or very similar, appear in biodata and reference measures--a suspicion the substance of which we categorically deny!

At any event, even the writer will concede that differential subgroup behavior in field study and experimental settings constitutes more convincing evidence of the behavioral significance of biodata subset membership than that already educed. Accordingly, in the academic field study context, selected subgroups have been shown to differ, in accordance with hypothesis, along the following lines: (a) in number of items marked "?" on a Thurstone scale of attitude toward the Negro--a sort of "band width" phenomenon negatively related to F authoritarianism; and (b) in test-taking motivation as measured by differential retest improvements in score on a scholastic aptitude test in response to a monetary incentive. In addition, Schoenfeldt (1970) has employed all the subgroups and has demonstrated that they differ significantly in academic performance, as evidenced by honors, dropouts and probations, and in the numbers of extracurricular activities in which they engage. Similarly, in the experimental domain, subgroups selected to address an hypothesis have been observed to differ in the meaningfulness of words, as derived from a verbal learning paradigm; in persuasability, as measured in a pretest-test design involving five types of persuasive appeals; in consistency of risk-taking behavior as measured by a modified poker dice game; and in general social desirability, which serves as a moderator of the tendency for perceived interpersonal similarity and attraction to be greater within than between subsets.

To date, the writer is aware of only two industrial studies utilizing, the procedures herein outlined: namely, those of Taylor and Ruda. Taylor

(1968), through the courtesy of Mr. Paul Sparks at Humble Oil & Refining, obtained the responses of 444 male, managerial employees to the Richardson, Bellows and Henry biodata form B. Using essentially the present methodology, he found that 84% of his salesmen in a primary sample of 222 came from 3 of 9 subgroups. He then employed a minimum distance classifier to assign the remaining 222 to the subgroups of the primary sample. Within this cross validation sample he found 83% of his salesmen in the 3 parallel subgroups.

Ruda (1970) had 458 executives of a large midwestern oil company complete a 247-item personal history form. Top level executives also ranked their subordinates' performances. All Ss were subgrouped employing the methods previously described and 13 subgroups were identified. Distance from superiors to subordinates subgroup was found to be negatively related to rated performance of subordinate. In addition, this distance measure of rater vs. ratee similarity was found to moderate relationships between biodata dimensions themselves and rated subordinate performance.

What important facts emerge from this brief review? First, persons who have behaved similarly in the past do tend to exhibit both similar concurrent and future behaviors. Second, confidence in the post hoc validity of biodata based subgrouping does seem justified; indeed, we lack any completely negative evidence to date. It thus seems reasonable to assume, following Taylor and Ruda, that subgroup membership will argue for a wide variety of differential behaviors of industrial moment and significance.

In delineating the relationship of this discussion to assessment, the reader will anticipate the writer. If we know that meaningful, distinctive, differential behavior is associated with subgroup membership, we may then identify an individual to be assessed, match his biodata profile with the subgroup profile it most closely resembles and predict the modal or typical behavior of the subgroup for the individual. To the

extent that the subset mean differs from the grand mean, and that the subset variance is less than the grand variance, the potential range of assessment errors is reduced. The efficiency of the procedure for the assessee is great, since he need only complete a biodata form. Its meaningfulness, on the other hand, is commensurate with our knowledge of the repertoire of subgroup behaviors.

Clearly, if we would implement this conceptual model in the industrial context we must not only administer an appropriate biodata form to large numbers of employees, but we must also be ingenious in the identification and development of large numbers of criteria of performance. Some of these will have immediate job relevance, some will have relevance in the future and some will seem to be only descriptors. But comprehensive description favors understanding and, with understanding, apparently irrelevant variables may turn out to represent critical moderators. The topic was treated more eloquently five years ago by then retiring president, S. Rains Wallace (1965), who noted that the answer to the question, "Criteria for what?" must always include "for understanding." Suffice it to say, in the present context, that the real vehicle of implementation would have to be a comprehensive system for the recording of <u>all sorts</u> of employee behavior. What does it mean to say that John Jones belongs to a given subgroup? It means as much as our knowledge of the behavioral characteristics of that subgroup! It is this knowledge we must have!

Clearly, it is easier to ask for more and different criteria than it is to identify some. What sorts of variables beyond the usual ratings of present performance and promotability might we consider. First, ratings of congeniality with associates would be useful, since persons within a subset tend to be attracted to each other, especially if the subset is a socially desirable one, and would presumably benefit from a chance to work together. Second, measures of "motivation", in contexts which range from

speed of performance to tenure, have strong subset affinity and obvious potential utility. Third, however appraised, managerial style in general and risk-taking behavior in particular, have clear affinity for subset and warrant further study in this context. Fourth, we have highly tentative data to suggest that innovativeness and ingenuity have subset affinity and should be re-examined accordingly. Fifth, let us take a purely hypothetical case and assume that the number of unexcused absences on ones record may be regarded as an inverse measure of his "responsibility." I would venture that such a variable would have both subset affinity and industrial relevance. In any case, it is illustrative of a most important point. No such predictor is needed to justify our collection of criterion data on responsibility. If responsibility has affinity for subset, and we can predict subset, we can indirectly predict not only responsibility but a host of other previously unpredictable variables as well. Incidentally, with testing somewhat under a cloud, it may be interesting to note that a full spectrum of test scores is predictable from biodata, and that criteria predictable from tests should therefore also be predictable with greater or lesser success from such biodata.

To the extent that the present conceptual model is a departure from traditional approaches, it has a new set of strengths and limitations. Some we can anticipate and some we will come upon as we proceed. One possible weakness is that we do not know precisely how to evaluate the match of individual to subset. In our Georgia study we did not subgroup some 25% of Ss at all because, within error limits, they fit two subsets equally well; or they were, conversely, outlyers who did not closely resemble any subset. Perhaps these subjects parallel those found in any testing program who have very average scores and about whom

decisions are made on other bases. In any event, as a second concern, some of our friends have argued that our model deals in foreordination and suggests that a man is completely a creature of his past. We, of course, have replied that we neither make nor enforce the rules, but only enunciate them. In any event, studies with employed persons suggest similar but slightly more numerous factors and subsets than we identify with college freshmen. Some continued but congruent differentiation with the years does thus seem to be implied. Third, in interpretation, both criterion and subgroup base rates are of obvious importance. If 84% of Taylor's (1968) entire sample had been salesmen instead of 31%, his findings would have seemed much less remarkable.

In a more positive vein, it is enormously efficient measurement if the completion of a biodata form, and the profiling of a S on it, can argue for that S's probable pattern of test scores, motivations, managerial style, promotability, creativity, and, indeed, much of his behavior repertoire. Comparably, and as priorly noted, any characteristics which adhere to subgroup membership may be inferred from it, whether they are directly measurable or not. Choice behaviors, illustrated by the sort of car one buys, provide a hypothetical example. Or further, since the observation of subgroup affinity for a given criterion is a post hoc observation, such a figure as the percent of Ss in a particular subset who satisfy a criterion of creativity may be expected to show sampling variation but not systematic skrinkage. The estimate provided is, in this sense, an unbiased one. Then, too, the longer records are kept, and the more differential information is attributed to subgroups, the more valuable the system becomes.

In conclusion, I have tried to sketch the broad outlines of a quantitatively based rationale and procedure for the comprehensive assessment of individuals. The overlap of subgroups with criteria is in the best

actuarial tradition, and may be regarded as the counterpart of more classical multiple regression approaches. What I believe to be somewhat more unique is the potential for attributing to the individual many behavioral characteristics of his subgroup which are neither directly measurable nor of immediate criterion relevance. At this point, the proposed model seems to speak directly to a concern expressed many times and in many different ways by sensitive students of personality. Some years ago, Allport (1937) feared that we were losing the individual through the interstices of nomothetic law. More recently Tyler (1959) has suggested that choice behavior might constitute the mortar to be used in bonding the concepts of trait theory into a workable psychology of individuality. In the end, many readers will recall a statement by Toops (1959) to the effect that the most meaningful thing one can say about an individual is that he belongs to a subset, or "ulstrith" the behavioral characteristics of which are known. Toops (1959) followed with a plea that we get about the behavioral description and definition of the ulstriths. Improvements in computer technology, statistical methodology and the sophistication of conceptual models have followed, and biodata subgroups are more sharply and complexly defined than ulstriths. The plea, however, still rings true, so let us <u>indeed</u> be about it, for now we have both a compelling rationale and a modus operandi!

References

Allport, G. W. Personality, New York, Holt, 1937.

Fiske, D. W. & Pearson, Pamela H. Theory and techniques of personality measurement. In P. H. Mussen & M. R. Rosenweig (Ed's). Annu. Rev. of Psychol. Palo Alto, Cal.: Annual Reviews, Inc. 1970.

Kelly, E. L. & Fiske, D. W. The prediction of success in the V. A. training program in clinical psychology. American Psychologist, 1950, 5, 395-406.

Owens, W. A. Toward one discipline of scientific psychology. American Psychologist, 1968, 23, 782-785.

Ruda, E. S. The effect of interpersonal similarity on management performance. (Doctoral dissertation, Purdue University, Lafayette, Ind., 1970)

Schoenfeldt, L. F. Life experience as a moderator in the prediction of educational criteria. Paper read at the meeting of the American Educational Research Association, Minneapolis, March, 1970.

Schoenfeldt, L. F. Nontest variables in the prediction of educational progress. In E. E. Dudek (Chm.), use of nontest variables in admission, selection, and classification operations. Symposium presented at the meeting of the American Psychological Association, Miami, September 1970.

Taylor, L. R. A Quasi-actuarial approach to assessment (Doctoral dissertation, Purdue University, Lafayette, Ind., 1968)

Toops, H. A. A research utopia in industrial psychology. Personnel Psychol., 1959, 12, 189-225.

Tyler, Leona. Toward a workable psychology of individuality. American Psychologist, 1959, 14, 75-81.

Wallace, S. R. Criteria for what? American Psychologist, 1965, 20, 411-417.

Ward, J. H. & Hook, M. E. Application of an hierarchical grouping procedure to a problem of grouping profiles. Educ. psychol. Measmt, 1963, 23, 69-81.